# Exploring social-indexical knowledge: A long past but a short history

PAUL FOULKES

*University of York*

*Abstract*

*This article has three main aims. First, I review what is known about encoding social-indexical information in phonetics|phonology, suggesting that it is more complex and more extensive than typically acknowledged in laboratory phonology. Secondly, I explore evidence for the development of indexical knowledge in phonological acquisition, considering indexical learning from the perspective of exemplar theory. One of the attractions of exemplar theory is its capacity to predict learning of both linguistic and social structures through the same mechanism. The tenets of exemplar theory also make a number of predictions about indexical learning: some types of indexical properties are likely to be transmitted more readily than others, as a function of the frequency and phonetic transparency with which they are manifested. The predictions appear to be borne out relatively well by the available data. I end by raising a number of implications and challenges, in particular, but not exclusively, for exemplar theory.*

## 1. Introduction

Language functions in a social context. All languages vary systematically depending on the communicative situation, and in correlation with the social backgrounds of speaker-listeners. Such facts have been acknowledged since the earliest attempts to theorise about language, dating back at least to Panini around 600 BC (Chambers 2002). The ubiquity and importance of systematic variation were recognised by the neogrammarian and structuralist schools, to which modern linguistics owes a great deal. Saussure, for example, stresses that language is fundamentally a system rooted in social context: "language never exists apart from the social fact, for it is a semiological phenomenon. Its social nature is one of its inner characteristics" (Saussure 1974 [1916]: 77). For theoretical reasons, however, the dominant models of linguistic theory in

JLP 1 (2010), 5–39
DOI 10.1515/LABPHON.2010.003

1836-6346/10/01–5

the twentieth century steered away from social matters, relegating issues of variation from serious consideration.

One of the defining themes of linguistic research over the last few years has been the revival of a healthy Kuhnian view of the interdependence of data and theory, and a reawakening of interest in "performance" data. A consensus has emerged to the effect that a cognitively-realistic, integrated theory of phonological knowledge, speech production, and speech perception must include more than an account of those properties pertaining to lexical contrast (e.g. Docherty 2007). It is clear that what speaker-listeners know about language involves not only abstract symbolic representations of "purely linguistic" structures, but also an extensive repository of social-indexical information. The latter includes knowledge about how to decode and interpret indexical information when it is encountered in listening, and how to encode it within speech production to signal aspects of personal identity and to achieve pragmatic goals in the course of speaking. Moreover, there is a growing body of evidence that indexical knowledge and linguistic knowledge interact (e.g. Nygaard et al. 1994; Goldinger 1997; Strand 1999; Hawkins and Smith 2001).

A central plank in the case for this consensus has been the widespread acceptance in cognitive science more generally of some role for episodic or exemplar theories of memory, coupled with usage-based statistical learning (Chater et al. 2006). Within linguistics support for this position has come from a remarkably broad spectrum of sources. This includes researchers in speech perception and psycholinguistics (e.g. Pisoni 1997; Luce et al. 2003; Chater and Manning 2006), phonetics (Pierrehumbert 2003; Johnson 2006), sociolinguistics (Foulkes and Docherty 2006; Docherty 2007), phonology (Bybee 2001; Becker et al. 2008; Coetzee and Pater in press), syntax (Jackendoff 2007), L1 acquisition (Abbot-Smith and Tomasello 2006; Vihman and Croft 2007; Beckman et al. 2007), L2 acquisition (Cutler and Weber 2007), and pragmatics (Plug 2006).

It now seems uncontroversial to conclude that social information is retained in memory alongside linguistic knowledge. Questions remain, however, over what sorts of social information are learned and stored, where and how they are stored in relation to linguistic information, and how social information affects linguistic processing. In considering these issues, I have three main aims in this chapter. First, I offer a review of what we know about social-indexical information at the level of phonetics/phonology, drawing on evidence from a range of disciplines. I suggest that social-indexical information is more complex and more extensive than is typically acknowledged in laboratory phonology. In turn this points to the necessity of sharing elements of methodology and theorising across disciplines. Secondly, I explore some of the predictions made by exemplar theory for the development of indexical knowledge. Exemplar theory predicts that some aspects of indexical learning should emerge earlier than others on account of factors such as frequency of exposure

and transparency of phonetic cues. Finally, I raise a number of implications and challenges for exemplar theory. These issues remain hypotheses to be tested, but which I think are important and potentially profitable lines of future work.

## 2. Indexical information in speech

### 2.1 *A working definition of indexicality*

In simple terms, indexical features in speech are those aspects of linguistic structure which correlate with non-linguistic factors (Abercrombie 1967). These factors include cross-speaker differences in gender, age, socio-economic status, ethnicity, group affiliations, regional background, and individual identity. They also include within-speaker effects related, for example, to emotion, attitude, and a broad range of context-specific factors which we can abbreviate as "style" (see further below; Foulkes 2006; Foulkes and Docherty 2006).

While this working definition of indexicality will suffice for present purposes, it should be borne in mind that indexicality has been the subject of considerable theorising within sociolinguistics and anthropology. The relationship between linguistic form and social category is not necessarily neat or easily defined (Ochs 1991; Silverstein 2003; Eckert 2008; and see Section 2.6 below).

### 2.2 *Biological and social roots*

Indexical features may be biologically- or socially-determined. Biological sources are involuntary, reflecting aspects of a speaker's anatomy and physiology, including effects of health. Socially-determined features are learned, and thus to some extent voluntary and controllable.

Although it is in principle possible to identify both biological and social sources of indexical information, in practice it may be rather difficult to distinguish them in a given sample of speech. It is apparent that gross biological differences, such as those between adult men and women, do not fully account for observed phonetic differences. For example, Johnson (2006) reports a cross-linguistic comparison of vowel formant values, showing that male-female differences vary considerably as a function of language.

Another example of the interaction between biological and social constraints is illustrated in Figure 1, which shows distributions of glottal variants of word-final pre-consonantal /t/ in a corpus of speech from Newcastle upon Tyne. The data were gathered to test a prediction that glottal variants should be disfavoured before voiceless fricatives (e.g. *hat shop*) as a consequence of incompatible demands on airflow (Pierrehumbert 1995). Voiceless fricatives require high airflow to generate turbulent noise, but the act of constricting the

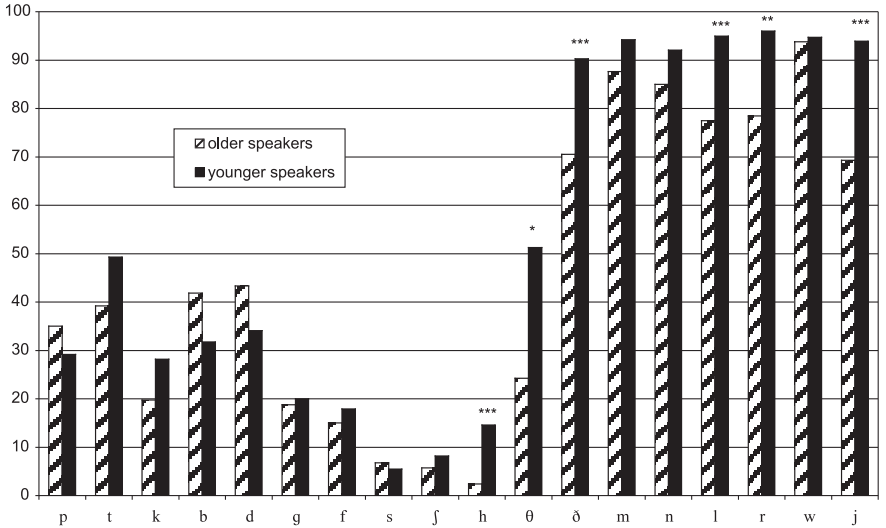


Figure 1. *Glottalisation rates for pre-consonantal /t/ in Newcastle English by age group. (* p < .05, ** p < .01, *** p < .001; N tokens = 4,883)*

glottis for [ʔ] interrupts the flow of air into the oral tract. Figure 1 offers strong support to Pierrehumbert's prediction. Glottal usage was indeed much lower before voiceless fricatives than other consonant categories. Slightly higher rates were observed before stops, which are also sounds that require relatively high airflow to generate adequate pressure for plosion. The highest glottal usage occurred before sounds that do not require high airflow: approximants, nasals, and liquids (and also /ð/, which is typically realised as an approximant). However, although the general patterns were consistent across generations, younger speakers (aged 15–27) used significantly more glottal stops than older speakers (45–67) in several consonantal contexts (indicated in Figure 1 by asterisks). The age-correlated differences are indicative of change in progress in the dialect, testifying to the interaction between socially-contingent constraints and those imposed by speakers' vocal tracts. A reflex of local learning is thus overlaid on the foundation provided by biology.

The inextricable linkage between the social and biological is perhaps clearest of all when we consider the properties of individual voices. As is well known from the literature in speaker identification, what makes a voice distinctive and recognisable is not a single feature or even a set of features that are permanently present (Nolan 1997). Rather, voices are highly plastic. Individuality is indexed by a complex set of interacting features which may be variable within certain parameters, and which may or may not be present in the voice at a given moment. This range of features comprises elements of biological endowment and learned aspects reflecting the social background of

the speaker and the environment in which speech is taking place. This very complexity is what makes speaker identification so difficult, and in part explains why we all make occasional mistakes in identifying familiar voices (Ladefoged and Ladefoged 1980, Foulkes and Barron 2000).

In sum, then, while biology certainly *constrains* variation, it does not determine it in isolation from social influences. Any speech act involves the simultaneous signalling of indexical information in parallel with propositional linguistic information, and it seems likely that the range of factors indexed will always include social ones (minimally reflexes of the individual speaker, and almost certainly of their sex and/or age). The relative contributions of the biological and social may vary, depending on the parameter under investigation. Vocal tract constraints provide most of the information necessary for a listener to detect a speaker's sex, while identifying regional origins depends rather more on locally-acquired pronunciation patterns.

### 2.3 *Quantitative and gradient nature of indexical features*

It is possible to find examples whereby indexical differences are expressed through usage of categorically different linguistic structures. For instance, speaker sex may be indicated by different verb or noun forms (Meyerhoff 2006: 203). However, such examples appear relatively rare, especially where phonetic/phonological features are concerned. More generally, indexical information is observable in quantitative differences. That is, a given feature may be used statistically more frequently by a particular social group than another. Alternatively, social differences may correlate with phonetic features occupying different (perhaps overlapping) positions along gradient scales. Where biology makes a large contribution to the indexical property it is possible that the cues are more transparent than in cases with a strong social component. For example, average fundamental frequency is largely dependent on biological factors. Although there is some overlap between f0 distributions for males and females, female f0 is on average 1.7 times that of male f0 (Klatt and Klatt 1990). On the other hand, social class differences in vowel production may be expressed via subtler statistical differences along arbitrary dimensions. For example, New Zealand English /e/ varies from [e] to [i] (Maclagan and Hay 2007). Younger speakers and women use closer variants than older speakers and men, and thus closeness can be said to index both sex/gender and age. The overlap between groups is much more marked than that for f0.

### 2.4 *Structural locations of indexical information*

In sociolinguistic studies of speech production, correlations have been observed between linguistic variants and social categories at many levels of structure. There is abundant evidence of within- and cross-speaker variation in

segmental structures, suprasegmental features (e.g. rhythm, intonation, voice quality, vocal setting, tonal alignment), and at a sub-segmental level (e.g. in gestural timing and magnitude) (Foulkes and Docherty 2006; Foulkes et al. 2010).

Perceptual evidence indicates that listeners can recognise indexical associations at various levels, using phonetic information to identify social information about the speaker or speech context. Thomas (2002) offers a summary of relevant studies. As an illustration, Foulkes et al. (in press) presented listeners with samples of Newcastle children's speech, and asked them to identify the talkers' sex. The focus of interest was whether listeners could use sociolinguistic variants as a cue. Previous studies of Newcastle adults had established strong gender patterning with respect to several variables (Foulkes and Docherty 2006). A characteristic feature of Newcastle English is the pronunciation of word-medial intersonorant /p t k/ (*happy, water, banker*). The local variants are fully voiced and lenited stops with laryngealisation. These variants have the status of local stereotypes, being above the level of conscious awareness (Labov 1994: 78). They attract explicit commentary by both insiders and outsiders. Plain oral stops also occur, but are mostly restricted to female speakers in more formal registers. Foulkes et al. (in press) presented three listener groups—one from the local Tyneside area and two control groups—with a series of mainly single-word examples extracted from recordings of children's speech. Responses were complex, with significant effects for a number of phonetic features, including amplitude, articulation rate and voice quality. However, with respect to the sociolinguistic variants, local listeners responded differently from the control groups. The results are summarised in Figure 2. The local listeners identified the talkers as girls significantly more frequently than the control groups did for word-medial tokens containing plain oral variants. This suggests that the local listeners were indeed able to use the variants as a cue to talker sex, presumably reflecting knowledge of the indexical values of these variants in the community at large. The control groups were not able to do so, since they do not share this indexical knowledge. Comparable results were found with pre-aspirated variants of stops in pre-pausal position.

Numerous sociolinguistic studies report associations between indexical categories and discrete linguistic structures such as individual consonants and vowels. This may be taken to imply that indexical knowledge is itself primarily attached to those abstract structures. It is clear, though, that indexical information may be imputed at other levels and types of structure. For example, Labov (2006) argues that social evaluation is primarily oriented to the lexicon rather than phonological structures. Indexical inferences may further be derived from simultaneous interpretation of many different components of a voice, as seems to be the case in speaker identification. Listeners may respond to combinations of segmental, suprasegmental, lexical, grammatical and prag-

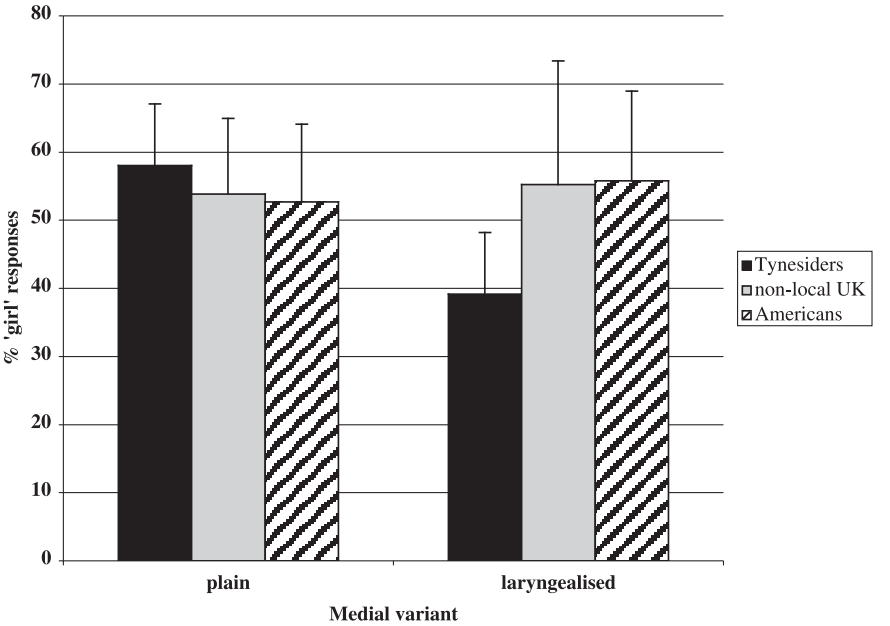


Figure 2. *% 'girl' responses to word-medial tokens by listener group and variant (Foulkes et al in press)*

matic features in identifying a talker (Rose 2002). There may therefore be an abstracted general model of a talker's voice available to listeners. Features lying outside what we generally assume to be linguistic knowledge may contribute as well: sounds such as coughs can be used by listeners to identify the person who produced them, suggesting that listeners may draw upon meta-level information about an individual's vocal profile in interpreting indexical information (Yarmey 2004).

While some aspects of indexical knowledge appear therefore to be represented in long-term memory, others are short-term and computed online, often in collaboration with interlocutors. For example, work in conversation analysis shows that phonetic details are used by participants in discourse to manage interaction, for example through negotiation of conversational turns by monitoring timing and f0 (Ogden 1999; Schegloff 2000; Bell et al. 2003; Local 2003).

## 2.5 *Indexical knowledge and phonological knowledge*

Given that a great deal of indexical information reflects learned properties, it follows that there must be some cognitive representation of that information. Anything that involves learning about indexical properties of sound

production, perception or patterning constitutes part of phonological knowledge, if we define the latter in the broad sense as "knowledge about sound" (Docherty and Foulkes 2000). Indexical knowledge permits a speaker to signal voluntary non-linguistic features through the medium of speech, and permits a listener to interpret the indexical values of voluntary or involuntary features. A key question is how that indexical knowledge is acquired, stored and processed relative to reflexes of linguistic knowledge (see further Sections 2.7 and 3 below).

### 2.6 *The extent of indexical variation*

The foregoing delimitation of indexicality is necessary because definitions, where given at all in experimental studies such as those in laboratory phonology, are often vague. Moreover, different disciplines have tended to focus on particular types of indexical information to the exclusion of other types. Some factors, notably sex, age, dialect, and individual voice, have figured reasonably prominently in phonetics and laboratory phonology. Others, however, have barely figured at all, leaving us with a restricted view of their contribution to phonological knowledge, speech production or perception. Group affiliation is a case in point. Sociolinguistic studies carried out in the frameworks of social networks and communities of practice indicate that linguistic variation depends to a large extent on voluntary or imposed group membership (Milroy 1987; Eckert and McConnell-Ginet 1999). Speech communities are not "pre-defined entities waiting to be researched" (Patrick 2002: 593). Instead, linguistic choices can be seen as symbolic capital, resources that can be drawn upon to establish personal and collegiate identity. Bucholtz (1999), for instance, points to the indexical value of various phonetic features in establishing "nerd" identity amongst a group of Californian students. The self-identified nerds typically used fully-released final [-t] instead of the unreleased or glottalised variants predominant in their peer group, and they did not participate in ongoing vowel changes such as fronting of /u:/. Similar highly-specific indexical patterns are reported by Kiesling (1998) for fraternity members, Munson (2007) for gay men, and Mendoza-Denton (2007a) for Latina gang members.

Furthermore, a great deal of sociolinguistic work has been devoted to deconstructing the broad labels typically referred to in discussing indexicality. The outcome of such work is a general scepticism about the explanatory value of coarse demographic categories such as those often employed in phonetics and laboratory phonology. These can be criticised as analyst constructs imposed on the data (Docherty 2007). It has been shown, for example, that variation linked to age and ethnicity can be better explained and interpreted if these factors are defined in socio-cultural rather than biological terms (Eckert 1997; Fought 2002). Class and style have likewise been shown to be convenient

heuristic labels rather than natural or inscrutable categories (Ash 2002; Schilling-Estes 2002). Stylistic variation, for instance, covers a very wide range of systematic effects related to formality, topic, interlocutor, the physical setting in which speech takes place, and the pragmatic exigencies of conversational interaction (Bell 1984).

The performed role of gender has also been shown to exert significant effects on linguistic choice, and may make a more important contribution in some circumstances than the binary biological category of sex (Cheshire 2002). A clear illustration is provided by Stuart-Smith (2007a), who discusses /s/ production by a socially heterogeneous group of 32 speakers from Glasgow. Although gross anatomical differences predict broad acoustic differences between males and females, the findings did not pattern along binary lines. Instead, the working class girls patterned more closely with the men in the sample, using a retracted articulation characterised by a relatively low centre of gravity in the fricative spectrum. The explanation for this is that social differences *between* females are extremely important in the local social context. Put simply, middle class and working class girls orient their identities with respect (or, rather, lack of it) to each other. While the working class girls employ the local dialectal variant of /s/, the middle class girls orient away from this because of its social-indexical saliency. Instead they adopt a much more sibilant realisation as one means to assert a particular gender identity, and to maintain distance from the social group they evaluate negatively.

The importance of social negotiation of indexical meaning is clearest of all when we examine sociolinguistic variation in communities where some form of abrupt social change takes place. This includes situations of language or dialect contact. Dyer (2002), for example, examined phonological variation in the English town of Corby. The town saw a large influx of steel workers from Scotland in the 1960s, creating a situation of dialect mixture. Scottish features have largely disappeared from the speech of subsequent generations, as is predicted by theoretical accounts of dialect contact. Generally speaking, minority variants disappear from the dialect mixture as a consequence of accommodation between speakers, a process which establishes and reinforces shared norms within the community (see Trudgill 1986, and Trudgill, Gordon, Lewis and Maclagan 2000 for a discussion of the contribution of this process to the development of New Zealand English). In the case of Corby, some of the originally Scottish features have been retained by subsequent generations, including [oː] in words of the GOAT lexical set. However, these no longer convey any indexical meaning pertaining to 'Scottishness'. Instead, they are recognised as markers of local identity, serving to distinguish Corby residents from inhabitants of neighbouring towns in a context of intense local rivalry. Heselwood and McChrystal (2000) provide a comparable example from Bradford, a city with a large Pakistani heritage community. Second and subsequent generations of migrant families retain phonological markers of ethnic identity in

their L1 English which were, for the first generation migrants who spoke English as L2, interference features from Punjabi. These include retroflex variants of /t, d/, non-velarised coda /l/, and deaspiration of initial voiceless stops. Moreover, the ethnic markers appear more prevalent in the speech of males.

The Corby and Bradford examples show that indexical values of linguistic features may be fluid and changeable, potentially subject to redefinition or renegotiation if the social context is conducive. They further demonstrate that indexical values may be the product of agency on the part of speaker-listeners, rather than the automatic reflexes of preordained categories. A consequence of this is that we should be wary of reifying those social categories typically addressed in (socio)linguistic research, or assuming uncritically either that they are objective, or even that they represent psychologically-real categories for the speaker-listeners under investigation (cf. comments by Pierrehumbert 2006). A further implication is that we should pay more attention to the performance of individuals in speaking and listening tasks in order to understand how indexical information is accessed. Although rarely an explicit focus of experimental research, exploring variation in individual results provides plenty of evidence that different people make different associations between phonetic forms and social categories, or that associations vary in robustness. Studies in forensic phonetics are instructive in this regard, since understanding individual variation may be a key concern in legal cases. They routinely show considerable variation in task performance across listeners and talkers (e.g. Foulkes and Barron 2000; Blatchford and Foulkes 2006). Another recent example is provided by Johnstone and Kiesling (2008), who explored perceptual responses among Pittsburghers to monophthongal realisations of /aw/. This feature has the status of a local stereotype, as evidenced by its orthographic encoding on tourist-oriented goods such as t-shirts (*down town* is rendered as <dahn tahn>). In spite of its apparent salience, they found that local listeners did not in fact agree on its indexical meaning. Some did not identify it as a local marker, while those who did so typically did not show high usage of the monophthong themselves.

## 2.7 *Summary*

Indexical variation is ubiquitous, pervasive, plastic, multifaceted, and complex. We incorporate indexical information whenever we speak, and encounter it whenever we listen to speech. Indexical information is packaged with linguistic information at every turn, and is often embedded in the same phonetic substance. It seems incontrovertible that there must be cognitive representation of indexical information, and that there must be communication between that part of indexical knowledge which pertains to phonetics and phonology, and phonological knowledge in the traditional linguistic sense. A mature system of speech perception and production must enable the speaker-

listener to encode and decode linguistic and indexical information through the same medium. From the perspective of a child learning language, a key task must be to identify which aspects of the input material convey linguistic information and which convey indexical information, such that appropriate linguistic and indexical structures may be extracted. This cannot always be a simple task, especially when the same phonetic material simultaneously encodes both types of information (Docherty et al. 2006).

However, it would be fair to say that most phonological models have little to say with respect to how the cognitive system handles the dual task of conveying and decoding indexical as well as linguistic information. Likewise, relatively little research has been carried out to explore how knowledge of indexical variation develops during the course of language acquisition. To a large extent this situation reflects the prevailing theoretical and methodological concerns of twentieth century linguistics. Work in developmental phonology has generally focused on the acquisition of language-specific targets, with much less interest in within-language or within-speaker variation (MacNeilage 1980). Moreover, while variation is the staple fare of sociolinguists, they in turn have paid little attention to children's speech. We therefore have a relatively poor understanding of what is learned, when and how it is learned, and of what relationship may obtain between indexical and linguistic knowledge.

In the next section I offer a review of the evidence for variation in the input which children receive, and which forms an essential component of language acquisition. I then examine the evidence from the perspective of exemplar theory. One of the attractions of exemplar theory is its capacity to predict and model learning of linguistic and non-linguistic structures through the same mechanism. The foundation of a store of detailed exemplars allows for generalisations of any type, in principle, to emerge through statistical learning. Provided relevant exemplars are present, associations can be made between forms and meanings, whether propositional or social (Smith and Zárate 1992). However, the central tenets of exemplar theory make a number of predictions about indexical learning. For instance, we can predict that some types of indexical properties are likely to be transmitted more readily than others, as a function of the frequency with which they are manifested. I explore some of these predictions, and assess how well they are borne out by the limited evidence currently available.

## 3. Acquisition of indexical information

Although there has been relatively little systematic study of variation in children's speech, it is clear from informal observation that children learn features specific to their local dialects, and they do so from a very early age. That children's speech should reflect the ambient input is unsurprising. Evidence of detailed learning reflecting local dialect input is clearly provided by e.g.

Kerswill (1996), Roberts (1997), and Imaizumi et al. (1999). However, regional dialect is just one of many dimensions of variability in the input available to children. Before assessing in more detail how children acquire indexical information, I first review what is known about variation in linguistic input.

### 3.1 *Indexical variation in input*

Under normal circumstances, children hear many different talkers. Their linguistic experience thus contains speaker-specific variability. Children are also exposed to the sorts of variation that are endemic to any speech community, reflecting sociolinguistic differences across speakers, and within-speaker variation, for example related to attitude, emotion, and a wide range of styles. A particularly clear example of social conditioning of variation can be found in bilingual or multilingual communities (which, we should remember, are in the majority on a global scale). The use and non-use of particular languages may be conditioned by a complex set of social rules. Depending on the family, languages may be associated with particular speakers or contexts of use. There may also be further dimensions of systematic variability, in that children face the task of learning the indexical values of variation within more than one language. They may hear also both L1 and L2 versions of their target languages (Khattab 2007).

In considering how indexical information is acquired, however, it is important to bear in mind that a child's primary input differs in its linguistic and phonetic composition relative to mature speech between adults. Several studies have noted broad prosodic differences between child-directed speech (CDS) and adult-directed speech (ADS). For example, CDS tends to be slower, and speakers use a higher overall f0 and wider f0 ranges (Fernald et al. 1989; Snow 1995; note that CDS and its properties are culture-conditioned, Ochs and Schieffelin 1995). Segmental production may also differ radically in CDS relative to ADS. Foulkes et al. (2005) examined consonant production in a corpus of CDS recordings with children aged 2 to 4 years from a working-class neighbourhood in Newcastle. In Section 2.4 above I outlined the phonetic and social patterns observed for intersonorant /p t k/. In conversational ADS, local laryngealised variants accounted for around 90% of all tokens in word-internal contexts (Foulkes and Docherty 2006: 414). Patterns in CDS, however, were quite different. The local variants occurred at an overall rate of just 36%, with a preference instead for plain oral stops. Individual mothers were variable in their use of the local variants in CDS, occupying positions across the entire continuum from 0 to 100%. Variation was also observed within CDS as a function of child age, with the use of standard variants decreasing across the age range. The concomitant increase in local variants appears to be symptomatic of a gradual shift towards a more adult-like interac-

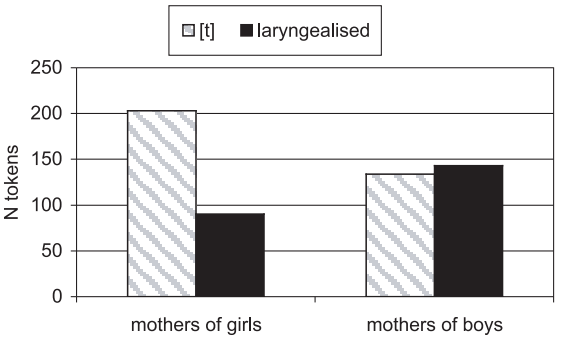


Figure 3. *Variant use for medial /t/ in CDS, by sex of child (Foulkes and Docherty 2006)*

tional style. The most striking difference of all, however, was in respect of child gender, with mothers of girls using far more standard forms in speech to their daughters than did mothers of boys in speech to their sons (Figure 3). Similar variation was observed in respect of /p t k/ in word-final contexts (e.g. *shut up*).

A study of variant patterns for stops in pre-pausal contexts revealed similar within-community variation (Foulkes and Docherty 2006). Pre-aspirated [ʰp ʰt ʰk] are innovative variants in Newcastle English, found most frequently in the speech of young women. Usage of pre-aspiration by the mothers in the CDS study varied from 0 to 89%, and this exact same range was found the children's own speech (see further section 4.3).

These patterns of variation have been mirrored in a study of CDS in Buckie, Scotland (Smith et al. 2007). Local dialectal forms in Buckie differ in many ways from standard English, although most speakers display variation between the vernacular and standard Scottish English. For example, the (au) variable (in words such as *house*) alternates between standard [ʌʉ] and vernacular [uː]. Usage of local forms varied substantially across the speaker sample, but there was a significant correlation between mother-child pairs. Smith et al. also explored stylistic variation within their corpus to test predictions that stylistic variation is transmitted to children on a dimension of formality (Labov 2001: 437). They classified exchanges in CDS, following a fairly simple categorisation system in line with Labov (2001). Formal activities were identified as those involving a main focus on teaching or discipline, while informal ones involved play or daily routine. Their findings support Labov's predictions, with local variants significantly more frequent in play (78%) and routine (68%) compared with discipline (31%) and teaching (19%). By contrast, a greater number of standard variants were used in the more formal settings.

It is important to note that the (au) variable in Buckie is, like the Newcastle voiceless stop variants, a local stereotype, above the level of social

awareness. By contrast, Smith et al. found no stylistic variation with grammatical variables which fall below the level of social awareness. The same is true for other phonological variables in Newcastle. Realisations of (ng) were examined in the CDS recordings, to test effects for style, following the method adopted by Smith et al. Variation between [ŋ] and [n] was expected. However, while this variable seems to elicit strong social evaluations in American studies (Labov et al 2006), the same does not appear true in the UK (Tagliamonte 2004). In the CDS corpus, use of non-standard [n] averaged 91%. Standard [ŋ] showed no significant correlation with style, and was mainly restricted to a few individual mothers and to *–thing* words.

In summary, examination of CDS data shows that children's input is not simply a scale model of the adult language. Nor is it indisputably a "simplified" version of the adult language, designed to assist children in identifying stable targets. From a purely statistical point of view, CDS may be even more variable than ADS: given the almost categorical usage of local variants of medial /p t k/ in ADS, a child in Newcastle would experience much more consistent input if paying attention to ADS than to CDS. To some extent CDS is its own tailored "style", diverging in its constitution from ADS. It also displays systematic internal variability, changing in line with children's age, and reflecting social differences which are relevant in the community. For example, the different CDS patterns presented to boys and girls in the Newcastle study presumably reflect gender-correlated differences in sociolinguistic norms: in simple terms, girls hear more standard forms than boys because there is a greater weight of social expectancy that they should grow up to use those forms (Foulkes et al. 2005).

### 3.2 *Predictions of an exemplar model of indexical learning*

Exemplar theory assumes that learned categories are substance-based, emerging from experience as the store of exemplars expands. As exemplars are added, language-learners appear to develop sensitivity to the statistical properties of the store. Exemplars with common properties begin to cluster together in memory, and category labels develop over statistical regularities. Pierrehumbert (2003) discusses how this process leads to the development of lexical and phonological categories. She further notes that sub-categories may emerge within the exemplar store contributing to a category. For example, while a cluster of phonetically-similar exemplars may yield a lexical category such as *cat*, distinct from other lexical categories, there may also be subordinate clusters of exemplars which share some but not all properties. Pierrehumbert offers the example of conditioned positional allophones, where, for example, instances of *cat* in turn-final position may end with released [tʰ], while others in pre-consonantal contexts may end in an unreleased [t˺] or glottal stop. A similar process to that described by Pierrehumbert may account for the learn-

ing of indexical patterns. In the case of a word like *water*, for example, a British English learner may experience exemplars which contain either [t] or [ʔ] as the medial consonant. The overall similarity of the two exemplar groups, reinforced by referential and top-down information, is enough for the two groups to cluster together, and for a single lexical representation to emerge from them. However, if the exemplars with [t] are distinguished from those with [ʔ] by virtue of social associations, then indexical knowledge may develop. For instance, if exemplars with [t] are predominantly heard from females while those with [ʔ] are generally from males, then knowledge of a gender-based correlation is predicted to emerge.

Following this line of reasoning, and taking stock of the range of indexical meanings signalled through speech, it is apparent that subcategories within an exemplar store will be the rule not the exception. We can expect there to be evidence of a potentially very wide range of indexical differences encoded within the exemplar store for all words and all language learners. Learners experience exemplars from many different talkers, from members of different social groups, and from different styles of speech. All exemplars will therefore contain indexical information, often in multiple layers. They thus provide the child with the raw materials for the acquisition of indexical categories as well as for the acquisition of abstract linguistic structures.

Any child's language development is a personal odyssey, dependent in part on the social and linguistic environment. Acquiring indexical knowledge is likely to be a long process–a life-long one if we include the learning of new voices, for example. Some properties may require longer to learn than others, and it is possible that some indexical meanings may never be identified at all (as noted by Clopper and Pisoni 2004 with reference to regional variation). The various categories of indexical knowledge are likely to emerge in overlap throughout the acquisition process as a child encounters new talkers and more and more examples of speech in social context.

However, exemplar theory makes some general predictions about the order of acquisition of indexical knowledge. Some categories of indexical information are likely to be both more transparently and more frequently represented than others, depending on their phonetic encoding in a growing exemplar store. The voice of the mother is experienced even *in utero*, and those of the main caregivers provide the majority of early input. Exemplars derived from these familiar talkers, incorporating indexical features particular to them, are likely to be high in frequency (leaving aside for the moment the issue of exemplar weighting). We might therefore expect to see clustering of exemplars as a function of individual talker in respect of those figures who are most influential in the child's world. We can also expect any child to hear speech from adult males, adult females, and other children. Talkers from these groups differ, of course, in relatively transparent, abrupt and consistent ways on account of major differences in vocal tract dimensions. Exemplars derived from

listening to males, females and children are therefore likely to partition in a relatively neat tripartite structure, as a direct consequence of major phonetic differences in f0 and vowel formant frequencies. Other indexical factors, by contrast, are less likely to provide robust cues within the exemplar store. These would include less familiar talkers, aspects of speech styles not typical in CDS or which are introduced later in the acquisition process, and aspects of regional or social variation not found in the immediate environment. Not only is there an obvious issue of frequency of exposure in such cases, but the relevant phonetic cues may also be less consistent or more arbitrary. Making inferences about associations between phonetic forms and indexical meanings is likely to be a more difficult job when those cues are infrequent, more variable in nature, and where the boundaries between categories may reflect small statistical differences in a particular parameter. Indexical knowledge of such factors is therefore predicted to require more time to emerge.

In sum, exemplar theory makes broad predictions about the order of indexical learning, based on (i) the overall contribution of indexical factors to the input, and (ii) the transparency of phonetic cues to the indexed category. Example predictions are summarised in Table 1.

Table 1. *Example predictions about the order of indexical learning.*

| | |
|---|---|
| earliest acquisition | maternal voice |
| early acquisition | other familiar individual talkers<br>biological sex<br>child versus adult<br>'style' features of early CDS |
| later acquisition | less familiar individual talkers<br>gender<br>'style' features beyond CDS |
| late acquisition | unfamiliar regional or social dialects |

### 3.3 *Indexical learning by children*

How well are these general predictions borne out by the available data? The answer seems to be "relatively well".

3.3.1 *Individual talkers* Learning of individual voices appears to develop immediately after birth. It has been demonstrated that neonates can identify the voice of their mothers by 2–4 days, showing that they can identify phonetic cues indexing a specific familiar talker (DeCasper and Fifer 1980). Indeed, it is possible that talker-specific information is learned *in utero*. Hepper

et al. (1993) played maternal speech to 36 week old foetuses through a loudspeaker placed on the mother's abdomen. They found evidence of discrimination between the mother's voice speaking normally and amplified. At 2 months infants responded to changes in talker when presented with a series of spoken stimuli (Miller 1983). ERP studies show that infants at 4 months process the mother's voice faster than other voices (Purhonen et al. 2005). Further evidence suggests that voice identification continues to develop through childhood. By 7–8 months infants can learn to identify new voices. In head-turning experiments, Houston and Jusczyk (2000, 2003) found that infants of 7.5 months showed longer orientation to passages containing familiar words when they had heard the talker previously. This suggests that at this age talker-specific information is retained by the infants, helping to prime their memory for words and assisting in word recognition. Comparable claims are made by Barker and Newman (2004), who found that familiarity with a voice helped infants at 7.5 months in separating two voices speaking simultaneously.

The effects found in such studies appear very short-term, however, and there is little if any evidence concerning long-term learning of voices by children. Few studies have been carried out, methodologies and tasks vary considerably, and it is very difficult to control for the many factors that can affect identification of voices (Bull and Clifford 1984). However, Bartholomeus (1973) studied nursery children aged of 4–5 years, and found that they could identify the voices of classmates almost as well as adults did. The best performers in fact bettered the results of the class teachers. Spence et al. (2002) found performance in identifying cartoon voices improved from 3 to 5 years, while in a cross-sectional study of voice identification, Mann et al. (1979) found children generally attained adult-like performance at age 10.

3.3.2 *"Style"* Style, as noted earlier, is an abbreviation for a very wide range of factors. Features of style are linguistic or phonetic features tailored to particular situations or registers. It appears that some reflexes of what we can categorise as style in this very broad sense may be learned even prior to birth. Hepper et al. (1993) found that neonates distinguished between "motherese" and normal speech. Both results indicate early sensitivity to broad prosodic differences, which in the former case are clearly indexical of the register adopted by mothers to infants.

Bilingual studies indicate that children learn socialised patterns of language, for example learning to use one language with particular interlocutors, by two years (Vihman 1985, Genesee 2006). In terms of within-language stylistic variation, production data suggest that social-indexical knowledge of stylistic difference may emerge gradually. In the Buckie study, children mirrored the stylistic variation of their mothers according to formality of interaction (Smith et al 2007). However, only the children aged 3 or over showed robust variation in their own speech. Gillen and Hall (2001) studied children at 3 years in

a game with a telephone box. Despite there being no-one at the other end of the toy telephone, some of the children displayed phonetic behaviour appropriate to telephone interaction, such as clearly defined speaking turns and stylised f0 contours in openings and closings. Andersen (1990) reports considerable stylistic variation in role playing games by children aged 4–7 years. All 24 of her subjects produced appropriate phonological and phonetic modifications. This included lowering f0 when talking in a "father" role, raising f0 in a "mother" role, and use of typical developmental segmental variants such as [f] for /θ/ and [w] for /r/. A five year old boy shifted to an apparently German accent in a "doctor" role, using creaky phonation and [v] for /w/. In similar studies, Sachs and Devin (1976) report that children as young as 2 years varied syntax and utterance length as a function of interlocutor, and Shatz and Gelman (1973) found similar interlocutor effects for 4 year olds, including in f0.

3.3.3 *Sex and gender* Infants as young as 4 months can match voices to faces of the same sex. Walker-Andrews et al. (1991) synchronised voices reading a two minute story with both male and female faces in a simultaneous presentation task. Children at 6 months showed preferential looking to the same-sex talker, while 4-month-olds showed the same effect on the second of two trials. A group of 3.5 month infants could not match for sex. (Patterson and Werker 2002 found sex matching for infants at 8 months but not 4.5 months. However, their tests involved isolated vowel sounds rather than extended speech samples.)

Miller's (1983) experiment cited in Section 3.3.1 revealed an interesting difference in performance by infants at different ages. Infants aged 2 months and 6 months were habituated to a group of voices from female or male speakers in a head-turn experiment. When presented with a new talker, 2-month-olds responded to the change irrespective of the sex of the talker, but 6-month-olds responded only to a change in sex category of the talker. This evidence might be taken to support the view that the younger children focus primarily on individual voices, while the older children have generalised to a more abstract category of talker, based on common phonetic properties such as f0. In a similar study, Miller et al. (1982) found that infants aged 7 months could discriminate between male and female voices when a series of stimuli spoken by talkers of one sex was interrupted by a talker of the opposite sex. These studies indicate knowledge of broad sex-based phonetic differences (rather than *gender*-based ones, as socially-defined).

From the perspective of speech production, a number of studies have examined f0, but with unclear or inconsistent results. Weinberg and Bennett (1971) report no statistical differences between boys and girls aged 5 and 6. Lee et al. (1995) found significant differences to emerge from age 12, and Perry et al. (2001) identified differences for 16 year olds but not for 4, 8 or 12 year olds.

Lieberman (1967) describes findings from a small-scale study, but his data are noteworthy. A 10-month-old boy was recorded with a mean f0 of 390 Hz when playing with his mother, but 340 Hz in a similar session with his father. A similar effect was found with a 13-month-old girl (average 390 Hz with the mother and 290 Hz with the father). The implication of these findings is that the children were adjusting their overall f0 level in relation to that of their interlocutor. It is unclear whether this is a reflex of accommodation to talker sex/gender, or to the individual talkers.

Studies of other phonetic parameters, however, have found robust differences between boys' and girls' speech. Perry et al. (2001) found systematic differences in formant values for four year olds. Some boys produced lower first and second formant values than girls, thus mirroring patterns found for adults. Children in Newcastle showed emergent gender-based differentiation in the use of pre-aspirated variants of voiceless stops, in line with patterns found in the adult community where pre-aspiration is strongly associated with young women (Foulkes and Docherty 2006). At ages 2;0, 2;6 and 3;0 there was no difference in usage of pre-aspiration between boys and girls, with the children as a group showing correlation with their mothers. However, the girls aged 3;6 used significantly more pre-aspiration than boys of the same age, and at 4;0 the same trend was observed. Roberts (1997) reports similar findings for /–t,d/ deletion in Philadelphia. Since these differences probably do not reflect physiological differences in the vocal tract at this age (although with respect to formant patterns there is some evidence that small physical differences in vocal tract anatomy emerge well before puberty: Crelin 1973), such differences are best interpreted as indicative of socially-conditioned gendered behaviour.

The age of 3 years seems to be an important milestone in children actively displaying gendered variation in speech production. Similar reports can be found on grammatical and discourse features (e.g. Gleason 1987; Ladegaard and Bleses 2003). It is unlikely to be a coincidence that this is also the age at which the child's awareness of gender emerges (Wells 1986; Bee 1998).

3.3.4 *Dialect* The ability to interpret different regional accents seems to emerge rather late, although there is remarkably little relevant research. Studies of children moving to new dialect areas shows that they are able to learn new phonological patterns, provided they move at an early enough age (under c. 11 years; Payne 1980; Chambers 1992). These findings indicate that the children must be able to perceive the accent-specific differences, although there is little evidence as to how quickly or easily the learning takes place.

The most significant study is that by Nathan et al. (1998), who compared 4 and 7 year old children's comprehension in tests with different regional accents of English. The children, who were from London, had to repeat and define words presented to them by speakers of London and Glasgow English. Predictably, they performed better in the same-dialect condition. The non-local

dialect proved difficult for many of the 4 year olds to comprehend, while in repetition tasks they often produced forms phonetically close, but not semantically equivalent, to the Glasgow forms. The difference in performance was much wider for children at 4 than at 7 years, suggesting that older children have developed a more robust perceptual system to deal with spoken variation. Le et al. (2007) show that effects of dialectal variation continue to be felt by adults in word recognition.

### 3.4 *Summary*

Although the available evidence is relatively scant, there does appear to be general support for the broad predictions of an exemplar-based model of indexical learning in which social categories are derived from statistical regularities in the input. Recognition of the maternal voice appears to be one of the earliest cognitive achievements. Identification of other voices emerges as a function of familiarity, and this is likely to be an ongoing process through the lifespan. The ability to generalise across male and female talkers also appears relatively early, as does an orientation to the general prosodic features of child-centred interaction. Evidence for recognition of more arbitrary indexical categories, with a strong socially-defined component, points to the age of around 3 years as a key stage. Tacit and explicit/overt *awareness* of the indexical information takes time to emerge, but variation in input provides children with the necessary raw materials from the start. There remains, of course, a great deal to be researched to explore further the development of indexical knowledge in a more systematic and consistent way than has hitherto been the case.

## 4. Implications and challenges

Exemplar-based learning of indexical categories receives general support, as we have seen in Section 3. Consideration of the evidence, however, also raises a number of interesting implications and challenges, some specifically for exemplar theory and some for phonetics and phonology more generally. I outline some of these briefly below. These issues are in many ways speculative, and constitute hypotheses to be tested. There are also issues which theoretical models, including exemplar theory, appear ill equipped to handle.

### 4.1 *Indexical learning assists language learning*

From the evidence discussed in section 3, it is clear that indexical learning from phonetic cues begins at or even before birth. Neonates can recognise the

maternal voice, and respond to prosodic cues in CDS. Clearly such knowledge precedes the formation of generalisations about the more abstract linguistic structure of the ambient language(s). We can infer, then, that early indexical learning provides a foundation for the learning of abstract linguistic structures. By virtue of being biased to the maternal voice through pre-natal experience, the child enters the world with an elementary framework for categorisation: the child is able to distinguish speech from other sounds, and the maternal voice from other spoken material. Similar conclusions have been drawn for a long time in the developmental literature, with reference to the emergence of abstract structures such as words and phonemes (e.g. Studdert-Kennedy 1983; DeCasper and Spence 1986; Houston and Jusczyk 2000, 2003). Early indexical categories may also serve as the basis for learning further indexical categories, for example as individual voices yield exemplars which cluster together via similar auditory or acoustic properties. Exemplar theory predicts the ways in which further categorisation (of linguistic as well as social material) may develop. However, its treatment of the initiation of the process does not consider either the potential contribution from the child's environment, or the possibility of innate language-specific endowments (see further Section 4.2 below).

A more general point can be made with respect to variation in acquisition. It is often assumed that variability in input poses a problem for the child learner, since, from the perspective of many theoretical models, the child's task is to derive stable target forms through analysis of a messy input. The messier the input, the harder the task is assumed to be. The characteristics of CDS as a register are typically portrayed as devices to render a "cleaner, simpler corpus from which to learn language" (Snow 1995: 180). However, variation may not be altogether problematic, or even dysfunctional. Obviously, if we take a broader view of what a child needs to acquire, to include knowledge of indexical information, then variation in the input is a pre-requisite to develop the knowledge necessary to produce and interpret socially-indexical variation.

### 4.2 *Parallel encoding of the indexical and linguistic*

Although there is considerable debate as to the precise nature of early representations, there is widespread support for the view that they are some way removed from the representations assumed for adults. Instead they appear surface-based, reflecting phonetic details of the input (with considerable debate as to precisely which details are attended to and represented; e.g. Jusczyk and Aslin 1995; Boersma and Levelt 2003; Vihman and Croft 2007). More abstract categorisation emerges towards the end of the single word period (c. 15–25 months, Vihman and Velleman 2000). The abstraction process is often countenanced within usage-based models, including exemplar theory. By implication, the details encoded within early representations are likely to

include reflexes of indexical information in respect of the talkers, dialects, and styles contributing to the input.

What remains unclear is what happens in later learning, once abstraction of linguistic structures begins. In some views, including many generative accounts, it is assumed that indexical details are eventually discarded altogether as irrelevant detail. This seems an unlikely scenario, however, in view of the compelling evidence that indexical details are present in long-term memory, and that they continue to affect perceptual processing. It is furthermore clear that retention of indexical information is vital for sociolinguistic competence.

One possibility is that phonological knowledge becomes gradually partitioned along the lines assumed in traditional models of grammar, such that the integrated information held in early word forms is gradually separated into two knowledge bases, one dealing with abstract information and a second with indexical information. Connections are likely to be retained between the two, rather than there being a total disjuncture (Docherty et al. 2006; Beckman et al. 2007). This might take the form of a scaffolding of abstract labels erected over the exemplar base, with connections retained to map between the two layers of representation (Pierrehumbert 2003). The degree of interaction between the two bases may vary according to learner, according to the situation, or both. Some tasks in speaking and listening may require careful attention to be paid to indexical details, for example, identifying the talker in a telephone call, or participating in conversations. Other tasks may permit or require a greater focus on more abstract structures, such as playing rhyming games and learning to associate spoken forms with orthographic ones. Some tasks in fact may require a constant shifting of attention from the abstract to the detailed, with linguistic and indexical information both present and both important, but to different degrees at a particular time. In a typical telephone call, for example, the opening exchange demands close attention to indexical information to identify the talker. The remainder of the call involves repeated shifts in importance for the abstract and indexical as propositional information is shared by the participants, and conversational turns are negotiated through appropriate cues.

Learning to balance attention to abstract and detailed information, no matter how these bodies of knowledge are stored cognitively, represents a challenge for a child acquiring language. We have as yet little idea of how such a balance is struck. Researchers in language development have tended to assume that a focus on abstract categories becomes the overriding concern. By contrast, the excellent body of research on the role of indexical information in linguistic processing has paid little attention to the variable nature of indexical information in respect of tasks, talkers or listeners (cf. Smith and Zárate's [1992] discussion of the shifting nature of social judgements). There is furthermore a dearth of perceptual work with child subjects, and relatively little focus on child data in exemplar theory.

While exemplar theory is well equipped to model the emergence of both linguistic and indexical categories, there is a lack of theoretical development regarding both the nature of abstraction, and constraints on the abstraction process. What kinds of linguistic primes and processes are derived from analysis of exemplars? What kinds of indexical categories can be posited? Are there constraints on the range of categories? How are linguistic and non-linguistic categories and processes stored and manipulated relative to one another? Work in speech perception has certainly considered the nature of perceptual units (Goldinger and Azuma 2003; Hawkins 2003, 2004), but researchers in the exemplar framework have generally avoided engagement with the discussions in mainstream phonological theory about, for example, whether the primes of lexical storage and processing are segmental or gestural, derivational or declarative (Cohn submitted). A notable counter-example is provided by Ettlinger (2007), who provides a computational simulation of chain shifting using an exemplar-based model. There has also been little engagement by proponents of exemplar theory with the universal grammar debate. Recent pro-UG work, however, has recognised the attraction of stochastic exemplar models, while concluding that UG is needed to constrain the learning process (Jackendoff 2007; Becker, Ketrez and Nevins 2008).

### 4.3 *Different children face different tasks in acquisition*

There is a further consequence of exemplar-based language acquisition. Learning is likely to differ from child to child. It is plain that every child experiences unique input. The exemplar store for any child will therefore itself be unique, and we can expect generalisations drawn from it to vary. To a large extent these variations in learning are likely to be trivial; given adequate input, similar generalisations can be made. However, there may be substantial and systematic differences across children, even in the same neighbourhood. Recall (Section 3.1) that children in the Newcastle study experienced statistically very different distributions of stop variants. There was variation on an individual basis and as a function of child age and sex. From a purely statistical point of view we might therefore expect to see reflexes of these substantial input differences in the children's output. That was indeed what was found on examination of the children's production. Pre-aspiration rates showed significant correlation between mothers and children in the group as a whole. There were also clear correlations in frequency for other variants, including for mother-child pairs in glottal stop usage for pre-pausal /t/ (a rare variant in this dialect) and for /t/ as [ɪ] in a restricted lexical set (Foulkes and Docherty 2006).

Assuming that the maternal model is the principal source of input, we can further hypothesise that early learning based on this input will by default include indexical information related to talker gender. Where there are gendered

sociolinguistic variables, for example, exemplars indexical of women will be present in abundance, and probably in far greater frequency than the alternatives indexing males. This presents a different challenge, then, for boys and girls in later acquisition. If early learning shows a strong link to the maternal model, boys may need to adjust certain phonetic parameters in their repertoire, either categorically or statistically, in order to conform to local gender norms. Boys may come to realise that the most frequent pattern in the exemplar store may not be the appropriate one for his developing gender identity. While girls can continue to follow the maternal model, including through use of its indexical properties, boys on the other hand must diverge from it. They need to identify gender-appropriate patterns, even if these are in low frequency in learned exemplars. A key stage in acquisition may therefore be a shift in attention to exemplars which may be in the minority in the exemplar store. Presumably this process is carried out in parallel with changes in social orientation. Exemplar theory needs to elaborate on how the shift takes place from emergence of categories through purely statistical processing to a system mediated by socially-weighted factors. Such issues are considered in more detail in the next section.

### 4.4 *Attention, saliency and exemplar weighting*

Frequency of exemplars plays an essential role in the development of abstractions across the exemplar store. However, the exemplars encoded in memory do not reflect raw experience, but rather, experience mediated by factors such as attention and saliency (Pierrehumbert 2006). What attention factors can we expect to shape a child's early learning of exemplars? Clearly, the person most important in a child's first few months of life is generally the primary caregiver. We might assume, then, that interaction with the primary caregiver has overwhelming importance. As a consequence, the exemplars derived from the input the caregiver provides should carry most weight, reinforcing the likely frequency effect resulting from the fact that interactions with the caregiver will outnumber those with other talkers. As life style changes in childhood, we can expect the importance of other talkers to increase. The peer group becomes increasingly important at the onset of schooling, for instance. Interaction with peers increases in frequency but also in social importance. We can thus predict that exemplars derived from peer interaction will take on greater importance and become more influential in the child's language development. This prediction appears to be borne out. Sudden and marked shifts in production norms are often reported, both anecdotally and in research findings, as children begin school (e.g. Kerswill and Williams 2000).

Mendoza-Denton (2007b) rightly points out that little work in the exemplar framework has explored the social mechanisms through which attention and saliency are mediated, or how weighting of exemplars for speech production

can be modelled. It is important to bear in mind also that speaker-listeners' indexical knowledge is often subconscious (Section 2.4). Thus, attention and saliency factors, and their effects on exemplar weighting, may differ according to whether indexical features are consciously or subconsciously accessed. Attention and saliency are also likely to vary across individuals and according to situation.

Given the great range of indexical factors inherent to speech, weighting is likely to be a complex and important process. Some assistance towards developing an adequate model of weighting is provided by sociolinguistic models. One prediction of accommodation theory, for example, is that face to face interaction outranks passive exposure, since it is generally more important in achieving communicative goals (Trudgill 1986). This is supported by many studies of child and adult interaction. In forensic speaker identification, for example, recall of a talker's voice appears better when the listener has had active exchange with the talker rather than passive exposure, suggesting more robust encoding of details about the voice (Hammersley and Read 1985). This is not to say, however, that memory is not affected by passive or non-interactive exposure. It is clear that children do pick up on details provided by spoken media. British parents are familiar with toddlers' penchant to sing pop songs in accents which resemble those of the original artist, and to parrot lines from American-accented toys. We have little idea of how such input is encoded in children's linguistic knowledge, however, as few have deemed it worthy of study (Stuart-Smith 2007b is a rare counter-example).

In addition, we should not neglect the role of the child himself as an agent in shaping his linguistic destiny. Production biases emerge in infants at around 9 months (McCune and Vihman 2001; Vihman and Croft 2007), as babbling preferences help to map adult input onto production templates. The child may also focus on his own output as a source of input, reinforcing the effects of the templates (Elbers 1997). Evidence for production templates is shown by systematic adaptation of adult forms to a common base. This occurs at a stage subsequent to initial word learning, where adult forms may be relatively well matched. For example, Vihman and Croft (2007) discuss adaptation of Estonian words by Vihman's daughter at 14–15 months. Her templates consisted of a limited consonantal inventory, all words were restricted to a single consonant type, and vowels were constrained such that open vowels preceded close ones, as illustrated in Table 2.

Production templates provide initial representational categories which may help shape the developing exemplar store for perception as well as production. Children as young as 10 months prefer listening to words which contain consonants similar to sounds used in babbling (Vihman and Velleman 2000). Thus early production helps to shape subsequent learning by focusing attention on subsets of the available input. There is a greater likelihood of extracting a robust exemplar if the input token conforms to the production template, while in

Table 2. *Adaptations of Estonian words by Virve Vihman (Vihman and Croft 2007)*

| Target | Adapted form(s) | Gloss |
|---|---|---|
| kallikalli | tati | hug |
| bravo | papu | bravo |
| lahti | ta \| ti | open |
| isa | asi | father |
| ema | ani, ami | mother |
| liha | ati | meat |

speech production greater importance is accorded to exemplars that are similar in form to the template. Weighting of exemplars may therefore commence through production biases, well before frequency effects begin to organise the exemplar store. Production templates can thus be said to focus attention and establish saliency for individual children.

### 4.5 *Early word forms may be far from citation forms*

Section 4.2 summarised the commonly held position that early word forms are surface-based and thus contain reflexes of indexical information. If this position is accepted, it follows that early representations may be some way distant from the representations commonly assumed in discussions of child (and, indeed, adult) phonology. Although it is rare to find explicit defence of such a position, most acquisition studies assume maximally full citation-like forms as target forms. Likewise, phonological treatments of adult language generally assume citation-like base forms. It is surely no coincidence that base/target forms are often indistinguishable from standard forms. However, for learners of non-standard dialects, citation forms may be difficult to access at all if evidence for them is scant within the input model. Yet citation forms are posited in most models as the base form, with processes, derivations, or constraint rankings organised to map to surface or output forms that may differ markedly from those base forms.

It is certainly true that reflexes of style variation are apparent in input from an early age, as we saw earlier. Such information provides evidence for the existence of abstract forms which may differ phonetically from the vernacular, and in many cases the more standard variant aligns closely with the citation form. However, such evidence may be limited to variables that are above the level of social awareness, such as glottal variants of /p t k/ in Newcastle and (au) in Buckie, where there is overt social recognition of the indexical values of the alternatives. Evidence for other citation features may not become apparent until much later, and may require support from orthographic learning (Port 2007). Hazan and Barrett (2000) report that children as old as 12 did not show phoneme categorisation skills comparable to those of adults. Studies of spell-

ing errors provide useful insight into children's cognitive organisation of sounds. Watson and Papen (2006), for example, collected spelling errors from 6 year olds in five British cities, to test the hypothesis that early phonological learning would interact with orthographic learning in different ways depending on the phonology of the local dialect. Accent-specific errors were indeed found. London children indicated vocalisation of coda /l/ (e.g. spellings of <aw, baw> for *all, ball*), but Newcastle children did not, as coda /l/ in this dialect is not velarised and thus does not vocalise. Newcastle children appeared to have no problem in rationalising initial and coda /l/ as a single category, whereas London children did.

In practice, the cognitive processing required of many children is likely to be the opposite of that assumed in typical phonological discussions. Rather than learning to map from citation base forms to variable surface forms, phonological knowledge must be built up to map *between* surface forms, or to map from them to citation forms. It is the latter that may require a long time to evolve, and which thus may have 'special' status, requiring the development of particular cognitive mechanisms. It is perhaps surprising that acquisition studies have not addressed empirically the issue of whether children have access to citation forms. It would be relatively simple to design appropriate studies, for example to compare whether children from different dialect backgrounds have equal access to citation forms in speaking or listening tasks. For the present I can offer only an anecdotal observation in support of the hypothesis that they probably do not. A teenager in Jane Stuart-Smith's Glasgow corpus comments as follows on relatives who have moved south: "m' Uncle . . . an' ma Auntie . . . live in London . . . s'awright, you can still underston' them".

## 5. Conclusions

Indexical variation in speech is exceptionally wide-ranging and complex. It is derived from many sources, both social and biological. Its constitution may vary in the course of speech acts and across speakers, and, for many reasons, listeners vary in their ability to identify and interpret it. The first and most obvious implication is that we should not limit our conception of indexicality, and subsequently our research designs, to a simple set of broad categories. What speakers and listeners learn, know and process is much richer than can be revealed by comparisons of males against females, or the effects of individual voices—defined as monolithic entities—can hope to reveal. Laboratory phonology, like phonetics and phonology more generally, has tended to focus on rather limited types of data. There are, of course, good reasons to experiment with isolated segments and syllables, nonsense words, and synthetic speech in an attempt to expunge indexical considerations from linguistic inquiry (Pierrehumbert et al. 2000). However, such materials permit insight into

a very limited slice of a speaker or listener's repertoire. The elimination of social context may, moreover, be an illusion. Engaging people in contrived tasks in an unnatural laboratory setting is such a special situation that it may not be appropriate to think of lab speech as 'socially neutral' (Huffman 2007: 115). While the effects of indexical information on perception and production may be suppressed, the context is so alien that we have little idea of its consequences for linguistic processing.

Extending experimental investigation further into the field is likely to be the most fruitful policy in the coming years of LabPhon. There is much to be learned from a deeper trawl into the literature in adjacent fields such as language acquisition, anthropology, dialectology, sociolinguistics, bilingualism, and conversation analysis, for insights into the social structures and indexical categories that have relevance for particular individuals in particular circumstances. In considering the extent and meanings of indexical features, and the role they play in linguistic processing, it is important to put *people* centre stage. This entails examining a wider range of forms of speech than is typically seen in laboratory phonology, including conversational interaction. It also suggests there may be much to learn from a thorough investigation of individual results in experimental research, in preference perhaps to the usual focus on group patterns (Docherty 2007, Scobbie 2007). Research is also needed on languages other than English, and in non-Western cultures. A final comment concerns the need to agree upon the definition of key terms, especially as more and more tributaries join the main flow of research in exemplar theory. Different subfields of language research have different understandings of terms such as frequency, variation, voice, lexicon, fine phonetic detail, and indexicality. What I have tried to do here is to delimit the latter to identify its ubiquity and pervasiveness, and thus to confirm its importance as a functional design feature of human speech.

While much remains to be done in understanding indexical knowledge, this is, to say the least, an interesting period in which to be working in phonology. Some have identified it as a new era, and a point of paradigm shift (e.g. Hawkins 2004; McLennan 2007). Exemplar theory appears to be the most promising candidate to construct a cognitively-realistic, integrated theory of phonological knowledge, speech production, and speech perception in which indexical knowledge is not marginalised but central. Exploring indexical knowledge may have a short history, but the future appears long and profitable.

**Acknowledgements**

I am grateful to the following for their time, thoughts and encouragement: Gerry Docherty, Bill Haddican, Jen Hay, Ghada Khattab, Jim Scobbie, Jane

Stuart-Smith, Marilyn Vihman, Dominic Watt, and Malcah Yaeger-Dror. Thanks also to the reviewers, Cynthia Clopper and Rob Podesva, for their numerous helpful suggestions. Finally, I record my thanks to the University of Canterbury's Erskine Fellowship programme, which provided support for the work discussed here and attendance at the LabPhon conference in Wellington.

Correspondence e-mail address: pf11@york.ac.uk

## References

Abbot-Smith, Kirsten & Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23: 275–290.

Abercrombie, David. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Andersen, Elaine S. 1990. *Speaking With Style: The Sociolinguistic Skills of Children*. London: Routledge.

Ash, Sharon. 2002. Social class. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 402–422.

Barker, Brittan A. & Rochelle S. Newman. 2004. Listen to your mother! The role of talker familiarity in infant streaming. *Cognition* 94: B45–B53.

Bartholomeus, Bonnie. 1973. Voice identification by nursery school children. *Canadian Journal of Psychology* 27: 464–472.

Becker, Michael, Nihan Ketrez & Andrew Nevins. 2008. The surfeit of the stimulus: analytic biases filter lexical statistics in Turkish devoicing neutralization. Web publication http://roa.rutgers.edu/view.php3?roa=1001

Beckman, Mary E., Benjamin Munson & Jan Edwards. 2007. Vocabulary growth and developmental expansion of types of phonological knowledge. In Jennifer Cole, & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 241–264.

Bee, Helen. 1998. *Lifespan Development* (2nd ed.). New York: Longman.

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Gregory, Michelle & Daniel Gildea. 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113: 1001–1024.

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13: 145–204.

Blatchford, Helen & Paul Foulkes. 2006. Identification of voices in shouting. *International Journal of Speech, Language and the Law* 13: 241–254.

Boersma, Paul & Claartje Levelt. 2003. Optimality theory and phonological acquisition. In Claartje Levelt, Lynn Santelmann, Maaike Verrips & Frank Wijnen (eds.), *Annual Review of Language Acquisition 3*. Amsterdam: John Benjamins, 1–50.

Bucholtz, Mary. 1999. 'Why be normal?': language and identity practices in a community of nerd girls. *Language in Society* 28: 203–223.

Bull, Ray & Brian Clifford. 1984. Earwitness voice recognition accuracy. In Gary L. Wells & Elizabeth F. Loftus (eds.) *Eyewitness Testimony: Psychological Perspectives*. Cambridge: Cambridge University Press, 92–123.

Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Chambers, Jack K. 1992. Dialect acquisition. *Language* 68: 673–705.

Chambers, Jack K. 2002. Studying language variation: an informal epistemology. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 3–14.

Chater, Nick & Christopher D. Manning. 2006. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences* 10: 335–344.

Chater, Nick, Joshua B. Tenenbaum & Alan Yuille. 2006. Probabilistic models of cognition: conceptual foundations. *Trends in Cognitive Sciences* 10: 287–291.

Cheshire, Jenny. 2002. Sex and gender in variationist research. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 423–443.

Clopper, Cynthia G. & David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32: 111–140.

Coetzee, Andries & Joe Pater. In press. The place of variation in phonological theory. To appear in John A. Goldsmith, Jason Riggle & Alan Yu (eds.) *Handbook of Phonological Theory* (2nd ed.). Oxford: Blackwell.

Cohn, Abigail C. submitted. Laboratory Phonology: past successes and current questions, challenges and goals. To appear in Cécile Fougeron & Mariapaola d'Imperio (eds.) *Laboratory Phonology 10*. Berlin: Mouton de Gruyter.

Crelin, Edmund S. 1973. *Functional Anatomy of the Newborn*. New Haven, CT: Yale University Press.

Cutler, Anne & Weber, Andrea. 2007. Listening experience and phonetic-to-lexical mapping in L2. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University, 43–48.

DeCasper, Anthony J. & William P. Fifer. 1980. Of human bonding: newborns prefer their mothers' voices. *Science* 208: 1174–1176.

DeCasper, Anthony J. & Melanie J. Spence. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9: 133–150.

Docherty, Gerard J. 2007. Speech in its natural habitat: accounting for social factors in phonetic variability. In Jennifer Cole & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 1–36.

Docherty, Gerard J. & Paul Foulkes. 2000. Speaker, speech, and knowledge of sounds. In Noël Burton-Roberts, Philip Carr and Gerard J. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, 105–129.

Docherty, Gerard J., Paul Foulkes, Jennifer Tillotson & Dominic J.L. Watt. 2006. On the scope of phonological learning: issues arising from socially structured variation. In Louis Goldstein, Douglas H. Whalen & Catherine T. Best (eds.) *Laboratory Phonology 8*. Berlin: Mouton de Gruyter, 393–421.

Dyer, Judy A. 2002. 'We all speak the same round here.' Dialect levelling in a Scottish English community. *Journal of Sociolinguistics* 6: 99–116.

Eckert, Penelope. 1997. Age as a sociolinguistic variable. In: Florian Coulmas (ed.) Handbook of Sociolinguistics. Oxford: Blackwell, 151–167.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453–476.

Eckert, Penelope & Sally McConnell-Ginet. 1999. New generalizations and explanations in language and gender research. *Language in Society* 28: 185–202.

Elbers, Loekie. 1997. Output as input: a constructivist hypothesis in language acquisition. *Archives de Psychologie* 65: 131–140.

Ettlinger, Marc. 2007. An exemplar-based model of chain shifts. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University, 685–688.

Fernald, Anne, Taeschner, T., Dunn, J., Mechthild Papoušek, Bénédicte de Boysson-Bardies & Fukui, I. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16: 477–501.

Fought, Carmen. 2002. Ethnicity. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 444–472.

Foulkes, Paul. 2006. Phonological variation: a global perspective. In Bas Aarts & April M.S. McMahon (eds.) *Handbook of English Linguistics*. Oxford: Blackwell, 625–669.

Foulkes, Paul & Anthony Barron. 2000. Telephone speaker recognition amongst members of a close social network. *Forensic Linguistics: The International Journal of Speech, Language and the Law* 7: 180–198.

Foulkes, Paul & Gerard J. Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34: 403–438.

Foulkes, Paul, Gerard J. Docherty & Dominic J.L. Watt. 2005. Phonological variation in child directed speech. *Language* 81: 177–206.

Foulkes, Paul, Gerard J. Docherty, Ghada Khattab & Malcah Yaeger-Dror. In press. Sound judgements: perception of indexical features in children's speech. To appear in Dennis Preston & Nancy Niedzielski (eds.) *A Reader in Sociophonetics*. Berlin: Mouton de Gruyter.

Foulkes, Paul, James M. Scobbie & Dominic J. L. Watt. 2010. Sociophonetics. To appear in William J. Hardcastle & John Laver (eds.) *Handbook of Phonetic Sciences* (2nd ed.). Oxford: Blackwell, 703–754.

Genesee, Fred H. 2006. Bilingual first language acquisition in perspective. In: Peggy McCardle (ed.), *Childhood Bilingualism*. Clevedon: Multilingual Matters, 45–67.

Gillen, Julia & Nigel Hall. 2001. 'Hiya, mum!': An analysis of pretence telephone play in a nursery setting. *Early Years* 21: 15–24.

Gleason, Jean B. 1987. Sex differences in parent-child interaction. In: Susan U. Philips, Susan Steele and Christine Tanz (eds.) *Language, Gender and Sex in Comparative Perspective.* Cambridge: Cambridge University Press, 189–199.

Goldinger, Stephen D. 1997. Words and voices: Perception and production in an episodic lexicon. In Keith Johnson & John W. Mullennix (eds.), *Talker Variability in Speech Processing*. San Diego: Academic Press, 33–66.

Goldinger, Stephen D. & Tamiko Azuma. 2003. Puzzle-solving science: the quixotic quest for units in speech perception. *Journal of Phonetics* 31: 305–320.

Hammersley, Richard & J. Don Read. 1985. The effect of participation in a conversation on recognition and identification of the speakers' voices. *Law and Human Behavior 9*: 71–81.

Hawkins, Sarah. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31: 373–405.

Hawkins, Sarah. 2004. Puzzles and patterns in 50 years of research on speech perception. In Janet Slifka, Sharon Manuel & Michelle Matthies (eds.) *From Sound to Sense: 50+ Years of Discoveries in Speech Communication,* Cambridge, MA: MIT. [http://www.rle.mit.edu/soundtosense/conference/pdfs/invitedspeakers/Hawkins%20PAPER.pdf]

Hawkins, Sarah & Rachel Smith. 2001. Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Rivista di Linguistica* 13: 99–188.

Hazan, Valerie & Sarah Barrett. 2000. The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics* 28: 377–396.

Hepper, Peter G., D. Scott & Sara Shahidullah. 1993. Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology* 11: 147–153.

Heselwood, Barry & Louise McChrystal. 2000. Gender, accent features and voicing in Panjabi-English bilingual children. *Leeds Working Papers in Linguistics and Phonetics* 8: 45–70.

Houston, Derek M. & Peter W. Jusczyk. 2000. The role of talker-specific information in word-segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 26: 1570–1582.

Houston, Derek M. & Peter W. Jusczyk. 2003. Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance* 29: 1143–1154.

Huffman, Marie K. 2007. Laboratory phonology and sociophonetics: partners in a conversation whose time has come. In Jennifer Cole & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 113–124.

Imaizumi, Satoshi, Kiyoko Fuwa & Hiroshi Hosoi. 1999. Development of adaptive phonetic gestures in children: Evidence from vowel devoicing in two different dialects of Japanese. *Journal of the Acoustical Society of America* 106: 1033–1044

Jackendoff, Ray. 2007. Linguistics in cognitive science: the state of the art. *The Linguistic Review* 24: 347–401.

Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: the emergence of social identity and phonology. *Journal of Phonetics* 34: 485–499.

Johnstone, Barbara & Scott F. Kiesling. 2008. Indexicality and experience: exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12: 5–33.

Jusczyk, Peter W. & Richard N. Aslin. 1995. Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29: 1–23.

Kerswill, Paul E. 1996. Children, adolescents and language change. *Language Variation and Change* 8: 177–202.

Kerswill, Paul E. & Ann Williams. 2000. Creating a new town *koine*: children and language change in Milton Keynes. *Language in Society* 29: 65–115.

Khattab, Ghada. 2007. Variation in vowel production by English–Arabic bilinguals. In Jennifer Cole & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 383–410.

Kiesling, Scott F. 1998. Variation and men's identity in a fraternity. *Journal of Sociolinguistics* 2: 69–100.

Klatt, Dennis H. & Laura C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87: 820–857.

Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal Factors*. Oxford: Blackwell.

Labov, William. 2001. *Principles of Linguistic Change. Vol. 2: Social Factors*. Oxford: Blackwell.

Labov, William. 2006. A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics* 34: 500–515.

Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski & Naomi Nagy. 2006. Listeners' sensitivity to the frequency of sociolinguistic variables. *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV34*. 12(2): 105–129.

Ladefoged, Peter & Jenny Ladefoged. 1980. The ability of listeners to identify voices. *UCLA Working Papers in Phonetics* 49: 43–51.

Ladegaard, Hans J. & Dorthe Bleses. 2003. Gender differences in young children's speech: the acquisition of sociolinguistic competence. *International Journal of Applied Linguistics* 13: 222–233.

Le, Jennifer T., Catherine T. Best, Michael D. Tyler & Christian Kroos. 2007. Effects of non-native dialects on spoken word recognition. Paper presented at Interspeech 2007, Antwerp.

Lee, A., Nigel Hewlett & Moray Nairn. 1995. Voice and gender in children. In Sarah Mills (ed.) *Language and Gender: Interdisciplinary Perspectives*. Harlow: Longman, 194–204.

Lieberman, Philip. 1967. *Intonation, Perception and Language*. Cambridge, MA: MIT Press.

Local, John K. 2003. Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics* 31: 321–339.

Luce, Paul A., Conor T. McLennan & Jan Charles-Luce. 2003. Abstractness and specificity in spoken word recognition: indexical and allophonic variability in long-term repetition priming. In Jeffrey S. Bowers & Chad J. Marsolek (eds.) *Rethinking Implicit Memory*. Oxford: Oxford University Press, 197–214.

Maclagan, Margaret & Jennifer Hay. 2007. Getting *fed* up with our *feet*: contrast maintenance and the New Zealand English 'short' front vowel shift. *Language Variation and Change* 19: 1–25.

MacNeilage, Peter F. 1980. The control of speech production. In Yeni-Komshian, Grace H., Kavanagh, James F. & Ferguson, Charles A. (eds.) *Child Phonology. Volume 1: Production*. New York: Academic Press. pp. 9–21.

Mann, V. A., R. Diamond & S. Carey. 1979. Development of voice recognition: parallels with face recognition. *Journal of Experimental Child Psychology* 27: 153–165.

McCune, Lorraine & Marilyn M. Vihman. 2001. Early phonetic and lexical development. *Journal of Speech, Language and Hearing Research 44:* 670–684.

McLennan, Conor T. 2007. Challenges facing a complementary-systems approach to abstract and episodic speech perception. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University, 67–70.

Mendoza-Denton, Norma. 2007a. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Oxford: Blackwell.

Mendoza-Denton, Norma. 2007b. Sociolinguistic extensions of exemplar theory. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 443–453.

Meyerhoff, Miriam. 2006. *Introducing Sociolinguistics*. London: Routledge.

Miller, Cynthia L. 1983. Developmental changes in male/female voice classification by infants. *Infant Behavior and Development* 6: 313–330.

Miller, Cynthia L., Barbara A. Younger & Philip A. Morse. 1982. The categorization of male and female voices in infancy. *Infant Behavior and Development* 5: 143–159.

Milroy, Lesley. 1987. *Language and Social Networks* (2nd ed.). Oxford: Blackwell.

Munson, Benjamin. 2007. The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech* 50: 125–142.

Nathan, Elizabeth, William H. G. Wells & Chris Donlan. 1998. Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language* 25: 343–365.

Nolan, Francis J. 1997. Speaker recognition and forensic phonetics. In William J. Hardcastle, & John Laver (eds.), *Handbook of Phonetic Sciences*. Oxford: Blackwell, 744–767.

Nygaard, Lynne C., Mitchell S. Sommers & David B. Pisoni. 1994. Speech perception as a talker-contingent process. Psychological Science 5: 42–46.

Ochs, Elinor. 1991. Indexing gender. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 335–358.

Ochs, Elinor & Bambi Schieffelin. 1995. The impact of language socialization on grammatical development. In Paul Fletcher & Brian MacWhinney (eds.), *Handbook of Child Language.* Oxford: Blackwell, 73–94.

Ogden, Richard A. 1999. A declarative account of strong and weak auxiliaries in English. *Phonology* 16: 55–92.

Patrick, Peter. 2002. The speech community. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 573–597.

Patterson, Michelle & Janet F. Werker. 2002. Infants' ability to match dynamic information in the face and voice. *Journal of Experimental Child Psychology* 81: 93–115.

Payne, Arvilla. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.) *Locating Language in Time and Space*. New York: Academic Press. pp. 143–178.

Perry, Theodore L., Ralph N. Ohde & Daniel H. Ashmead. 2001. The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America* 109: 2988–2998.

Pierrehumbert, Janet B. 1995. Prosodic effects on glottal allophones. In Osamu Fujimura & Minoru Hirano (eds.) *Vocal Fold Physiology*. San Diego: Singular Publishing Group, 39–60.

Pierrehumbert, Janet B. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–154.

Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34: 516–530.

Pierrehumbert, Janet B., Mary E. Beckman & D. Robert Ladd. 2000. Conceptual foundations of phonology as a laboratory science. In Noël Burton-Roberts, Philip Carr & Gerard J. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, 273–303.

Pisoni, David B. 1997. Some thoughts on 'normalization' in speech perception. In Keith Johnson & John W. Mullennix (eds.), Talker variability in speech processing. San Diego: Academic Press, 9–32.

Plug, Leendert. 2006. *Phonetic Reduction and Pragmatic Organisation in Dutch Conversation: a Usage-Based Account*. PhD dissertation, University of York.

Port, Robert. 2007. How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology* 25: 143–170.

Purhonen, Maija, Riitta Kilpeläinen-Lees, Minna Valkonen-Korhonen, Jari Karhu & Johannes Lehtonen. 2005. Four-month-old infants process own mother's voice faster than unfamiliar voices: electrical signs of sensitization in infant brain. *Cognitive Brain Research* 24: 627–633.

Roberts, Julie. 1997. Acquisition of variable rules: a study of (-t, d) deletion in preschool children. *Journal of Child Language* 24: 351–372.

Rose, Philip. 2002. *Forensic Speaker Identification*. London: Taylor & Francis.

Sachs, Jacqueline & Judith Devin. 1976. Young children's use of age-appropriate speech styles in social interaction and role-playing. *Journal of Child Language* 3: 81–98.

Saussure, Ferdinand de. 1974 [1916]. *A Course in General Linguistics*. London: Fontana.

Schegloff, Emmanuel A. 2000. Overlapping talk and the organization of turn-taking for conversation. *Language in Society* 29: 1–63.

Schilling-Estes, Natalie. 2002. Investigating stylistic variation. In: Jack K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.) *Handbook of Language Variation and Change*. Oxford: Blackwell, 375–401.

Scobbie, James M. 2007. Biological and social grounding of phonology: variation as a research tool. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University, 225–228

Shatz, Marilyn & Rochel Gelman. 1973. The development of communication skills: modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development* 38: 1–38.

Silverstein, Michael. 2003. Indexical order and the dialectics of social life. *Language and Communication* 23: 193–229.

Smith, Jennifer, Mercedes Durham & Liane Fortune. 2007. 'Mam, my trousers is fa'in doon!': Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19: 63–99.

Smith, Eliot & Michael Zárate. 1992. Exemplar-based model of social judgment. *Psychological Review* 99: 3–21.

Snow, Catherine E. 1995. Issues in the study of input: Finetuning, universality, individual and developmental differences, and necessary causes. In Paul Fletcher & Brian MacWhinney (eds.), *Handbook of Child Language*. Oxford: Blackwell, 180–193.

Spence, Melanie J., Pamela R. Rollins & Susan Jerger. 2002. Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research* 45: 214–222.

Strand, Elizabeth. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18: 86–99.

Stuart-Smith, Jane. 2007a. Empirical evidence for gendered speech production: /s/ in Glaswegian. In Jennifer Cole & Jose Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 65–86.

Stuart-Smith, Jane. 2007b. The influence of the media. In: Carmen Llamas, Louise Mullany & Peter Stockwell (eds.) *The Routledge Companion to Sociolinguistics*. London: Routledge, 140–148.

Studdert-Kennedy, Michael. 1983. On learning to speak. *Human Neurobiology* 2: 191–195.

Tagliamonte, Sali A. 2004. Someth[in]'s go[ing] on!: variable ing at ground zero. In Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fidell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundergren & Mats Thelander (eds.) *Language Variation in Europe: Papers from ICLAVE 2*. Uppsala: Uppsala Universitet, 390–403.

Thomas, Erik R. 2002. Sociophonetic approaches of speech perception experiments. *American Speech* 77: 115–147.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis & Margaret Maclagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36: 299–318.

Vihman, Marilyn M. 1985. Language differentiation by the bilingual infant. *Journal of Child Language* 9: 249–253.

Vihman, Marilyn M. & William Croft. 2007. Phonological development: toward a 'radical' templatic phonology. *Linguistics* 45: 683–725.

Vihman, Marilyn M. & Shelley Velleman. 2000. Phonetics and the origin of phonology. In Noël Burton-Roberts, Philip Carr & Gerard J. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, 305–339.

Walker-Andrews, Arlene S., Lorraine E. Bahrick, Stacy S. Raglioni & Isabel Diaz. 1991. Infants' bimodal perception of gender. *Ecological Psychology* 3: 55–75.

Watson, Kevin & Uta Papen. 2006. Speech and spelling: facilitation or interference? Paper presented at Sociolinguistics Symposium 16, University of Limerick. [http://www.lancs.ac.uk/staff/watsonkd/watsonss16ho.pdf]

Weinberg, Bernd & Suzanne Bennett. 1971. Speaker sex recognition of 5- and 6-year-old children's voices. *Journal of the Acoustical Society of America* 50: 1210–1213.

Wells, Gordon. 1986. *The Meaning Makers*. Portsmouth: Heinemann.

Yarmey, A. Daniel. 2004. Common-sense beliefs, recognition and the identification of familiar and unfamiliar speakers from verbal and non-linguistic vocalizations. *International Journal of Speech, Language and the Law* 11: 267–277.